Эксперименты ученых показали, что человеческое ухо более чувствительно к изменениям звука на низких частотах, чем на высоких. То есть, если частота звука изменится со 100 Гц на 120 Гц, человек с очень высокой вероятностью заметит это изменение. А вот если частота изменится с 10000 Гц на 10020 Гц, это изменение мы вряд ли сможем уловить.

психофизическая величина, подобранная "под человека".

В связи с этим была введена новая единица измерения высоты звука — мел. Она основана на психо-физиологическом восприятии звука человеком, и логарифмически зависит от частоты:

mel=1127.01048ln⁡(1+freq700)  
Собственно, мел-спектрограмма — это обычная спектрограмма, где частота выражена не в Гц, а в мелах. Переход к мелам осуществляется с помощью применения мел-фильтров к исходной спектрограмме. Мел-фильтры представляют из себя треугольные функции, равномерно распределенные на мел-шкале. В качестве примера здесь изображены 10 мел-фильтров (на практике их берут больше, здесь их мало для наглядности):

Резюмируя все вышесказанное: на мел-спектрограмме сохраняется больше информации, которая хорошо воспринимается и различается человеком, чем на обычной спектрограмме. Иными словами, такое представление звука больше сфокусировано на низких частотах, и меньше — на высоких.